



OIPE

RAW SEQUENCE LISTING

DATE: 02/04/2002

PATENT APPLICATION: US/09/978,242

/09/978,242 TIME: 11:38:59

Input Set : N:\Crf3\RULE60\09978242.raw
Output Set: N:\CRF3\02042002\1978242.raw

SEQUENCE LISTING

```
3 (1) GENERAL INFORMATION:
             (i) APPLICANT: Bandman, Olga
      5
                            Yue, Henry
      6
                            Corley, Neil C.
      7
                            Shah, Purvi
      8
            (ii) TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN
     10
           (iii) NUMBER OF SEQUENCES: 3
     12
            (iv) CORRESPONDENCE ADDRESS:
     14
                  (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
     15
     16
                  (B) STREET: 3174 Porter Drive
     17
                  (C) CITY: Palo Alto
                                                             ENTERED
                  (D) STATE: CA
     18
     19
                  (E) COUNTRY: USA
     20
                  (F) ZIP: 94304
     22
             (V) COMPUTER READABLE FORM:
     23
                  (A) MEDIUM TYPE: Diskette
     24
                  (B) COMPUTER: IBM Compatible
     25
                  (C) OPERATING SYSTEM: DOS
                  (D) SOFTWARE: FastSEQ for Windows Version 2.0
     26
     28
            (vi) CURRENT APPLICATION DATA:
                  (A) APPLICATION NUMBER: US/09/978,242
C--> 29
                  (B) FILING DATE: 15-Oct-2001
C--> 30
                  (C) CLASSIFICATION:
     31
     33
           (vii) PRIOR APPLICATION DATA:
                  (A) APPLICATION NUMBER: 09/241,333
     35
     36
                  (B) FILING DATE:
                  (A) APPLICATION NUMBER: 08/990,114
     38
     39
                  (B) FILING DATE:
          (viii) ATTORNEY/AGENT INFORMATION:
     41
     42
                  (A) NAME: Billings, Lucy J.
                  (B) REGISTRATION NUMBER: 36,749
     43
                  (C) REFERENCE/DOCKET NUMBER: PF-0451 US
     44
            (ix) TELECOMMUNICATION INFORMATION:
     46
                  (A) TELEPHONE: 650-855-0555
     47
                  (B) TELEFAX: 650-845-4166
     48
     49
                  (C) TELEX:
        (2) INFORMATION FOR SEQ ID NO: 1:
     52
     54
             (i) SEQUENCE CHARACTERISTICS:
     55
                  (A) LENGTH: 545 amino acids
     56
                  (B) TYPE: amino acid
     57
                  (C) STRANDEDNESS: single
```

(D) TOPOLOGY: linear

58

DATE: 02/04/2002 TIME: 11:38:59

PATENT APPLICATION: US/09/978,242

60	(vii)		IMMEDIATE SOURCE: (A) LIBRARY: TLYMNOT05													
61			(A)	LIB	RARY	: TL	OMMY	T05								
62			(B)	(B) CLONE: 2809795												
64			SEQU										_			
66		Ala	Thr	Glu		Val	Asn	Gly	Asn		Thr	Glu	Glu	Pro		Asp
67	1	1			5	- 1		G	a1	10	Dh.a	a 1	mb	Lou	15	A an
68	Thr	Thr	ser		vaı	тте	HIS	ser		ASII	Pne	GIII	THE	30	reu	ASP
69	. 1 -	a1	T	20	01 =	T	170]	ת 1 ת	25	T 110	Lou	Nan	Clu		Tur	Val
70 71	Ala	GIÀ	Leu 35	Pro	GIII	Lys	val	40	13 LU	гуѕ	ьец	ASP	45	iie	тут	Val
72	7 l a	C1 11	Leu	Val	λla	Иie	Sar		Len	Agn	Glu	Ara		Tle	Glu	Ala
73	ALG	50	LCu	V CI I	n Lu		55	ПОР	neu	p	014	60				
74	Leu		Glu	Phe	Asn	Glu	Asp	Gly	Ala	Leu	Ala	Val	Leu	Gln	Gln	Phe
75	65	-1-				70		1			75					80
76	Lys	Asp	Ser	Asp	Leu	Ser	His	Val	Gln	Asn	Lys	Ser	Ala	Phe	Leu	Cys
77	-	_			85					90					95	
78	Gly	Val	Met	Lys	Thr	Tyr	Arg	Gln	Arg	Glu	Lys	Gln	Gly	Thr	Lys	Val
79				100					105					110		
80	Ala	Asp	Ser	Ser	Lys	Gly	Pro		Glu	Ala	Lys	Ile		Ala	Leu	Leu
81			115					120		_,	1	~ 1	125	_		
82	Glu		Thr	Gly	Tyr			Asp	Val	Thr	Thr		GIn	Arg	гàг	туг
83	~ 1	130	_	ъ.	D		135	17- 7	m	C 0 m	C1	140	Cln	Dro	Cor	Wa l
84		GIY	Pro	Pro		150	ser	vai	туг	ser	155	GIII	GIII	PIO	ser	160
85 86	145	mbr	Glu	T10			Clv	Lvc	Τlρ	Pro		Δsp	Len	Phe	Glu	
87	сту	1111	GIU	116	165	vaı	Gry	гуз	TIC	170	лгу	пэр	LCu	1110	175	пър
88	Glu	Leu	Val	Pro		Phe	Glu	Lvs	Ala		Pro	Ile	Trp	Asp	Leu	Arg
89	314	200		180				_1 -	185				-	190		
90	Leu	Met	Met	Asp	Pro	Leu	Thr	Gly	Leu	Asn	Arg	Gly	Tyr	Ala	Phe	Val
91			195					200					205			
92	Thr	Phe	Cys	Thr	Lys	Glu	Ala	Ala	Gln	Glu	Ala	Val	Lys	Leu	Tyr	Asn
93		210					215					220				_
94		His	Glu	Ile	Arg		Gly	Lys	His	Ile		Val	Cys	Ile	Ser	
95	225					230			_	- 1	235	_		Ŧ	ml	240
96	Ala	Asn	Asn	Arg		Phe	Val	GIY	ser		Pro	Lys	ser	ьуs	255	гуs
97	C1	<i>a</i> 1 -	т1.	T a	245	Cl.	Dho	Cox	Tura	250	Thr	clu	Clu	Lou		λen
98 99	GIU	GIN	Ile	260	GIU	GIU	Pne	ser	265	Val	1111	GIU	GIY	270	1111	кър
100	Va l	т1/	2 T 01		uic	Cln	Dro	λer		1.77C	I.V.C	T.v.c	. Acr		r Gly	y Phe
101	٧aJ	L 114	275		піз	GII	. FIC	280		л Бүз	, шуз	Пус	285		, 01	, 1110
102	Cvs	s Phe			ı Tvr	Glu	Asp			Thr	Ala	Ala			a Arc	g Arg
103	011	290		. 010		010	295		-1-			300				, ,
104	Arc			Ser	Gly	Lys	Val	Lys	. Val	Trp	Gly	Asn	val	Gly	y Thi	r Val
105	305				_	310		_			315					320
106	Glu	ıTrı	o Ala	Asp	Pro	Ile	Glu	Asp	Pro	Asp	Pro	Glu	ı Val	L Met	t Ala	a Lys
107					325					330					335	
108	Va]	L Ly:	s Val			Val	Arg	Asn			Asn	Thr	Val			ı Glu
109				340					345		_	_		350		
110	11ϵ	e Lei	ı Glu	ı Lys	: Ala	Phe	Ser	Glr	ı Ph∈	e Gly	' Lys	Leu	ιGlu	ı Ar	g Val	l Lys

DATE: 02/04/2002 TIME: 11:38:59

PATENT APPLICATION: US/09/978,242

111		355					360					365				
111	Lys Leu		Nan	Tree	λla	Dho		Uic	Dho	λαn	clu		λen	Clv	λla	
112 113	175 Leu 370		АБР	туг	АТа	375	116	птэ	Pile	кър	380	ALG	АЗР	Gly	nia	
	Val Lys		Mot	C111	Clu		Acn	Clu	Tvc	7 cn		Clu	C137	Clu	Λen	
114		АТА	Met	GIU	390	мес	ASII	GIY	гуу	395	Leu	GIU	СТУ	GIU	400	
115	385	Tla	v. 1	nh a		T	Dwo	Dwo	700		Luc	λκα	Lvc	Clu		
116	Ile Glu	rre	val		Ald	гуѕ	PLO	PIO	_	GIII	цуъ	Arg	гуѕ	415	AIG	
117		G1 -	3	405	3 1 -	31.	T	3	410	Wat	П	Nan	Nan		Mrr.	
118	Lys Ala	GIII	-	GIII	Ald	Ald	гуѕ		GIII	мес	тут	ASP	430	1 7 1	тут	
119		G1	420	D == 0	rri a	Wa+	Dwo	425 Dmg	Dwo	mbr	A 350 GT	(1)		C1,,	7. r:c1	
120	Tyr Tyr		Pro	Pro	HIS	мес		PIO	Pro	1111	Arg	445	Arg	СТУ	Alg	
121	al., al.,	435	Q1	a 1		C1	440	Dwo	70 20 0	7.00	m		C1**	Marx	Clu	
122	Gly Gly	_	GIÀ	GTA	туг		гуг	PIO	PIO	ASP		1 7 1	GTÅ	тут	GIU	
123	450				.	455	m			TT 2 -	460	m	3	G1	.01	
124	Asp Tyr	Tyr	Asp	Tyr	_	GIA	Tyr	Asp	Tyr		ASII	тÀт	Arg	GIY		
125	465	_	_	_	470	a.1		a.1	_	475		77 - 7	a1	3. 1 -	480	
126	Tyr Glu	Asp	Pro		Tyr	GIY	Tyr	GIU		Pne	GIN	vaı	GIY		Arg	
127				485	- 1			~ 1	490		_	~		495		
128	Gly Arg	Gly	_	Arg	GLY	Ala	Arg		Ala	Ala	Pro	ser		GTÄ	Arg	
129			500	_	_		_	505	~ 3	,	_	~ 3	510	a l	21	
130	Gly Ala		Pro	Pro	Arg	Gly		Ala	Gly	Tyr	Ser		Arg	GIY	GIY	
131		515			- 1	1	520		- 1	_	_	525		a 1		
132	Pro Gly		Ala	Arg	Gly		Arg	Ala	GIY	Lys		GIY	Arg	GTÀ	Arg	
133	530					535					540					
134	Ser															
135	545															
137	(2) INFO				_											
139	(1)	SEQ														
140	· ·															
	41 (B) TYPE: nucleic acid 42 (C) STRANDEDNESS: single															
142								Le								
143		•) TO				ır									
145	(vii)						>m 0 E									
146		•) LII				7105									
147	,) CLO						NO.							
149	(X1) GGGCGCG	SEQ									.73330	amere i	ATTC (accani	ртесте	60
151			-1 1 -1 1	ACCGC	JG AC	3060	CICC	GAL		1010	CAAC	1001	AIC (120
150					70 00	20010	n monon an	C 7 7	CACC	amaa	A CITIC	CCCC	700 /	א גריי איר		120
152	CCAGCGG	CGT (GAGC:	r'rcg(100
153	CCAGCGG GCAGCGA	CGT (GAGC' CGCG'	TTCG(TTTC(CC GC	CAACO	CCGAT	ACC	CATCO	GGAC	AGG <i>A</i>	ATTT	CTC (CGCC:	TCAGCC	180
153 154	CCAGCGG GCAGCGA CAACGGG	CGT (GCA (GAG /	GAGC' CGCG' ATCT(TTCG(TTTC(CTGG <i>I</i>	CC GC	CAAC(CATG(CCGAT GCTAC	T ACC	CATCO	GAC GTT	AGGA AATO	ATTT(GGAA	CTC (CGCCT STACT	TCAGCC TGAAGA	240
153 154 155	CCAGCGG GCAGCGA CAACGGG GCCCATG	CGT (GCA (GAG A GAT A	GAGC' CGCG' ATCT(ACTA(TTCGC TTTCC CTGG <i>I</i> CTTC	CC GC AA AC FG C	CAACO CATGO AGTTA	CCGAT GCTAC ATCC	T ACC C AGA A TTC	CATCO AACAI CAGA <i>I</i>	GGAC FGTT AAAT	AGGA AATO TTTO	ATTT(GGAA CAGA(CTC (ATG (CAT 1	CGCC GTAC TGCT	TCAGCC TGAAGA TGATGC	240 300
153 154 155 156	CCAGCGG GCAGCGA CAACGGG GCCCATG TGGTTTA	CGT (GCA (GAG AGAT AGAT AGAT AGAT AGAT AGAT AGA	GAGC' CGCG' ATCT(ACTA(CAGAA	TTCGC TTTCC TTGGA CTTCT AAGTT	CC GC AA AC IG CA IG CC	CAACO CATGO AGTTA TGAAA	CCGAT GCTAC ATCCA AAACT	T ACC C AGA A TTC T AGA	CATCO AACAT CAGAA ATGAA	GGAC FGTT AAAT AATT	AGGA AATO TTTO TACO	ATTTO GGAA CAGAO GTTGO	CTC (ATG (CAT 1 CAG (CGCCT GTACT GGCTT GGCT	rcagcc rgaaga rgatgc Agttgc	240 300 360
153 154 155 156 157	CCAGCGG GCAGCGA CAACGGG GCCCATG TGGTTTA ACATAGT	CGT (GCA (GAT AGAT AGAT AGAT AGAT AGAT AGAT AGA	GAGC' CGCG' ATCT(ACTA(CAGA) TTAG)	TTCGC TTTCC TTGGA CTTCT AAGTT ATGAA	CC GC AA AC IG CA IG CO AA GA	CAACO CATGO AGTTA TGAAA AGCTA	CCGAT GCTAC ATCCA AAACT ATTGA	ACC AGA TTC AGA AGC	CATCO AACAT CAGAA ATGAA CTTTA	GGAC GTT AAAT AATT AAAA	AGGA AATC TTTC TACC GAAT	ATTTO GGAAA CAGAO GTTGO CTCAA	OTC (ATG (CAT (ATG (ATG (ATG (ATG (ATG (ATG (ATG (A	CGCCT GTACT GGCTT GGCTA AAGA(CAGCC GAAGA GATGC AGTTGC CGGTGC	240 300 360 420
153 154 155 156 157 158	CCAGCGG GCAGCGA CAACGGG GCCCATG TGGTTTA ACATAGT ATTGGCA	CGT (GCA (GAT (GAT (GAT (GAT (GAT (GAT (GAT (GA	GAGC: CGCG: ATCT(ACTA(CAGA) CTAGA	TTCGC TTTCC TTGGA TTCC AAGTC ATGAA AACAC	CC GC AA AC IG CA IG CC AA GA GT TC	CAACO CATGO AGTTA TGAAA AGCTA TAAAO	CCGAT GCTAC ATCCA AAACT ATTGA GACAC	ACC AGA TTC AGA AGC TGA	CATCO AACAI CAGAA ATGAA CTTTA ATCTO	GGAC GGTT AAAT AATT AAAA CTCT	AGGA AATO TTTO TACO GAAT CATO	ATTTO GGAA! CAGAO GTTGO FTCA! GTTC!	OTC O ATG O CAT 1 CAG O ATG A AGA A	CGCCT GTACT GGCTT GGCTA AGAGA	CAGCC CGAAGA CGATGC AGTTGC CGGTGC AAGTGC	240 300 360 420 480
153 154 155 156 157 158 159	CCAGCGG GCAGCGA CAACGGG GCCCATG TGGTTTA ACATAGT ATTGGCA CTTTTTA	CGT (GCA (GAT (GAT (GAT (GAT (GAT (GAT (GAT (GA	GAGCT CGCGT ATCTC ACTAC CAGAA CTAGA CTTCA GGAGT	TTCGC TTTCC TTGGA CTTCT AAGTT ATGAA AACAC	CC GC AA AC IG CA IG CA AA GA GT TA	CAACO CATGO AGTTA TGAAA AGCTA TAAAO GACTA	CCGAT GCTAC ATCCA AAACT ATTGA GACAC	ACC AGA TTC AGA AGC GTGA	CATCO AACAT CAGAA ATGAA CTTTA ATCTO AGAGA	GGAC GGTT AAAT AATT AAAA CTCT AGAA	AGGA AATO TACO GAAT CATO AAAO	ATTTO GGAAA CAGAO GTTGO GTTCAA CAAGO	OTC CATG CATG ATG AGA AGA AGGA GGA GGA	CGCCT GTACT GGCTA AGACA ACAAA	TCAGCC TGAAGA TGATGC AGTTGC CGGTGC AAGTGC	240 300 360 420 480 540
153 154 155 156 157 158 159 160	CCAGCGG GCAGCGA CAACGGG GCCCATG TGGTTTA ACATAGT ATTGGCA CTTTTTA	CGT (GCA (GAT / GAT / GAT / GAT (GAT (GAT (AGT / AGT /	GAGCT CGCGT ATCTC ACTAC CAGAL CTAGAL CTAGAL CTAGAL CTAGAL CTAGAL CTAGAL CTAGAL CTAGAL CAGAL CAGA	TTCG(TTTC(TTGG! TTTCT AAGTT AACA(FCAT(GACC!	CC GC AA AC IG CA IG CA AA GA GT TI GA AC AG AI	CAACO CATGO AGTTA IGAAA IAAAO GACTI IGAGO	CCGAT GCTAC ATCCA ATTGA GACAC GACAC GCAAA	T ACC C AGA A TTC T AGA A AGC G TGA G GCA A AAT	CATCO AACAI ATGAI ATCTO AGAGI	GGAC GGTT AAAT AAAA CTCT AGAA GGCA	AGGA AATO TACO GAATO AAAO CTCT	ATTTO GGAA CAGAO GTTGO GTTCA CAAGO	ETC CATG CATG ATG AGA AGA AGA CAAA CAAA C	CGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	TCAGCC TGAAGA TGATGC AGTTGC CGGTGC AAGTGC AAGTGC AAGTGC AAGTAGC AGGATA	240 300 360 420 480 540
153 154 155 156 157 158 159 160 161	CCAGCGG GCAGCGA CAACGGG GCCCATG TGGTTTA ACATAGT ATTGGCA CTTTTTA AGATTCT CACACTT	CGT (GCA (GAT (GAT (GAT (GAT (GAT (GAT (GAT (GA	GAGCT CGCGT ATCTC ACTAC CAGAL CTTAGL CTTCL GGAGT AAAGC GTGAC	ITCG(ITTCC ITGGA ITTCT AAGTT AAGA ICATC ICATC ICACC ICACC	CC GC AA AC IG CA AA GA AA GA AA AC AA AC AAG AC IG GA AG AC AG GA	CAACC CATGO AGTTA CGAAA CAAAC GACTI CGAGO ACAGA	CCGAT GCTAC ATCCA ATCGA ATTGA GACAC GCAA AGGAA	ACC AGA TTC AGA AGC TGA AGC AGC AGC AGC AGC AGC AGC AGC AGC A	CATCO NACAI CAGAI NTGAI NTCTO NGAGI NTGGI	GGAC GGTT AAAT AAAA CTCT AGAA GGCA AGGA	AGGA AATO TTTO TACO GAATO CATO AAAO CTCT CCAO	ATTTO GGAAL CAGAO GTTGO TTCAL GTTCL CAAGO TTGGL	CTC CATG CATG CATG AAGA AAAA CAAAA CAAAA CAAAA CAAAA	EGCCT EGCTT EGCTA AGAGA ACAAA ECAAA ECAAA	TCAGCC TGAAGA TGATGC AGTTGC CGGTGC AAGTGC AGTAGC AGTAGC AGTAGC AGGCTA CGGTTA	240 300 360 420 480 540 600 660
153 154 155 156 157 158 159 160 161 162	CCAGCGG GCAGCGA CAACGGG GCCCATG TGGTTTA ACATAGT ATTGGCA CTTTTTA AGATTCT CACACTT	CGT (GCA (GAT / GAT (GAT / GAT (CAG (CAG (CAG (CAG (CAG (CAG (CAG (CAG	GAGCT CGCGT ATCTC ACTAC CAGAL CTAGAL CTTCL GGAGT AAAGC GTGAC CAGCC	TTCGC TTTCC TTGGA TTCT AAGTT AAGAA AACAC TCATC GACCA CCACC CCACC	CC GC AA AC IG CA AA GA AA GA AA AC AAG AA IG GA AIG GA AIG GA AIG GA AIG TI AA GA	CAACC CATGC AGTTA TGAAA TAAAC GACTT TGAGC ACAGA	CCGAT GCTAC ATCCA ATTGA ATTGA GCACA GCACA AGGAA ACTGA	T ACC A TTC A AGC A AGC A AGC A AAT A GTA A GAT	CATCO NACAT CAGAL ATGAL ATGTO NGAGL ATGGL ATGGL ATGGL ATGGL	GGAC FGTT AAAT AAAT AAAA CTCT AGAA GGCA AGGA	AGGA AATO TTTO TACO GAATO AAAO CTCT CCAO GGAA	ATTTO GGAAL CAGAC GTTCAL GTTCL CAAGC TTGGA CCTCC	CTC CATG CATG AAGA AAGA AAAA CAAAA CAAG AAAA CAAAA CAAAA CAAAA CAAAA CAAAA CAAAAA CAAAAAA	EGCCT EGCTT EGCTT AAGAG ACAAA ECAAA EAACA ATTCG	TCAGCC TGAAGA TGATGC AGTTGC CGGTGC AAGTGC AGTAGC AGTAGC AGTAGC AGGCTA CGTTTA AGATCT	240 300 360 420 480 540 600 660 720
153 154 155 156 157 158 159 160 161	CCAGCGG GCAGCGA CAACGGG GCCCATG TGGTTTA ACATAGT ATTGGCA CTTTTTA AGATTCT CACACTT	CGT (GCA GAT GAT GAT GAT GAT GAT GAT GAT GAT GA	GAGCT CGCGT ATCTC ACTAC CAGAL TTAGL CTTCL GGAGT AAAGC CAGCC CACC CAGCC CACCC C	TTCGC TTTCC TTTCC TTGGA AAGT AATGAA AACAC TCATC GACC TCATC TTCT TTCT	CC GC AAA AC IG CA AAA GA GT TT GAAA AC AG	CAACC CATGC AGTTA CGAAAC FACTT CGAGCA ACAGA ATTAT	ECGAT GCTAC ATCCA ATTGA ATTGA GCACA GCACA ACTGA TTTGA	T ACC AGA A TTC AGA A AGC A AGC A AGC A AGC A AGC A GTA A GTA A GAA	CATCO AACAT ATGAA ATGTO ATGTO ATGGA ATGGA ATGGA AAGCT AAGCT	GGAC GGTT AAAT AATT AAAA CTCT AGAA GGCA AGGA GGTG	AGGA AATO TACO GAATO AAAO CTCT CCAO GGAA CCTA	ATTTO GGAAA CAGAC GTTGA GTTCA CAAGC TTGGA CCTCC AAGAT ATATC	CTC CATG CATG CATG AATG AATG AAAA CAAA C	EGCCT EGCTT EGCTT AAGAG ACAAA ECAAA ECAAA ATTCG EAAGA	TCAGCC TGAAGA TGATGC AGTTGC CGGTGC AAGTGC AGTAGC AGTAGC AGGCTA CGTTTA AGATCT	240 300 360 420 480 540 600 660

PATENT APPLICATION: US/09/978,242 TIME: 11:38:59

DATE: 02/04/2002

166	165	AGAA	GCA	GCT (CAGG	A/GGC'	rg T	TAAA	CTGT	A TA	ATAA	ГСАТ	GAA	ATTC	GTT (CTGGA	AAAACA	A 900	0
1680 TGAGATCAC AAAACASCG ATGACAAGAA AAAAACACA GGCTTTGCGT TTCTGAATA 1080	166	TATI	GGT	GTC '	TGCA'] `TC	AG T	TGCC/	AACA	A TAC	GGCT'	тттт	GTG	GGCT	CTA '	TTCC:	raaga(G 960	Э
1680 TGAGATCAC AAAACASCG ATGACAAGAA AAAAACACA GGCTTTGCGT TTCTGAATA 1080	167	TAAA	ACC	AAG (GAAC	AGAT'	TC T	TGAA	GAAT	TA(GCAA <i>i</i>	AGTA	ACA	GAGG	GTC '	TTACA	AGACG:	г 1020	0
140																			0
TATGGCANAG GTNANAGTGC TGTTTGTACG CAACCTTGCC AATACTGTAA CAGAAGAGTT 1260 1717 17	169																		0
TATGGCANAG GTNANAGTGC TGTTTGTACG CAACCTTGCC AATACTGTAA CAGAAGAGTT 1260 1717 17		CTGG	GGGZ	AAT	GTTG	GAAC'	rg T	TGAA'	TGGG	C TG	ATCC'	ГАТА	GAA	GATC	CTG I	ATCC:	rgagg:	г 1200	0
172 TITAGAAAAG GCATITAGTC AGITTIGGAA ACTGGAACGA GTGAAGAGAGT TAAAAGATTA 1320 13																			0
1380		TTTA	GAA	AAG (GCAT'	TTAG	TC A	GTTT	GGGA	A AC'	TGGA	ACGA	GTG	AAGA	AGT '	TAAA	AGATTA	A 1320	0
144		TGCG	STTC	ATT (CATT	T'TGA'	IG A	GCGA	GATG	G TG	CTGT	CAAG	GCT	ATGG	AAG	AAAT	GAATG	G 1380	0
156	174	CAAA	AGAC'	TTG	GAGG	GAGA	AA A	TATT	GAAA'	r TG	TTTT	rgcc	AAG	CCAC	CAG A	ΑΤΟΛ	CAAAA	3 1440	0
TTATGGTCCA CCTCATATGC CCCCTCCAAC AGGAGGTGGA GGGCGTGGAS GTAGAGGTGG 1560 177 TTATGGATAT CCTCCAGATT ATTATGGATA TGAAGATTAT TATGATTATT ATGGTTATGAA 1620	175																		О
171	176	TTAT	GGT	CCA (CCTC	ATAT	GC C	CCCT	CCAA	CAA	GAGG!	rcga	GGG	CGTG	GAG (GTAGA	AGGTGG	G 1560	Э
THACCATAAC TATCGTGGTG GATATGAAGA TCCATACTAT GGTTATGAAG ATTTTCAAGT 1680 1740 1760 17	177																		О
1740		TTAC	CATA	AAC '	TATC	GTGG	IG G	ATAT	GAAG	A TC	CATA	CTAT	GGT'	TATG	AAG A	TTTA	rcaag:	г 1680	О
1800 GGCTGCTCCT CCCCGCGGTA GAGCCGGTTA TTCACAGAGA GAGGTCCTG GATCAGCAAG 1800 1860 AGGCGTTCGA GCAGGGAAAA GGGGTCCGAG CCCGTCCTGA CCTGTACAA TAAGAGCATCA 1860		TGGA	AGCTA	AGA (GGAA	GGGG1	rg g	TAGA	GGAG	C AAG	GGGG'	TGCT	GCT	CCAT	CCA (GAGG1	CGTG	G 1740	О
1816																			О
182 CTTGCTATGT GGGATTACAC CAGAAGCTTG CAGTGGAGTA ATGGTAAGGA AATCAAGCAA 183 CCTTAAATAT GTCGGCTGTA TAGGAGCATA TTCTATTGCA GAAGACCTTC CTATGAAGAT 1980 1980 1981 CATGGAATCA AATACGGGAC ATTGAACTAA TACTTGGACT TTGATATGAA TTTCTTTAAC 2040 185 AATTTTCTCT GCAGTGCAAG TTATTAAACT AAAGCTACT 187 (2) INFORMATION FOR SEQ ID NO: 3: 189 (i) SEQUENCE CHARACTERISTICS: 190 (A) LENGTH: 714 amino acids 191 (B) TYPE: amino acid 192 (C) STRANDEDNESS: single 193 (D) TOPOLOGY: linear 195 (vii) IMMEDIATE SOURCE: 196 (A) LIBRARY: GenBank 197 (B) CLONE: 128842 199 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3: 201 Met Val Lys Leu Ala Lys Ala Gly Lys Thr His Gly Glu Ala Lys Lys 202 1 5 10 15 203 Met Ala Pro Pro Pro Lys Glu Val Glu Glu Asp Ser Glu Asp Glu Glu 204 20 25 30 205 Met Ser Glu Glu Glu Asp Asp Ser Ser Gly Glu Glu Val Val Ile Pro 206 35 40 45 207 Gln Lys Lys Gly Lys Lys Ala Thr Ala Thr Pro Ala Lys Lys Val Val 208 50 60 209 Val Ser Gln Thr Lys Lys Val Ala Val Pro Thr Pro Ala Lys Lys Ala 210 65 70 75 80 211 Ala Val Thr Pro Gly Lys Lys Ala Ala Ala Thr Pro Ala Lys Lys Ala 212 85 90 95 213 Val Thr Pro Ala Lys Ala Val Ala Thr Pro Gly Lys Lys Gly Ala Thr 214 100 105 110 215 Gln Ala Lys Ala Leu Val Ala Thr Pro Gly Lys Lys Gly Ala Thr 216 115 120 125 217 Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp Ser																			О
183																			Э
2040																			Э
185																			Э
187 (2) INFORMATION FOR SEQ ID NO: 3: 189		AATT	TTC	TCT (GCAG'	rgca/	AG T	TATT	AAAC	r AA	AGCT	ACT						2079	9
189 (i) SEQUENCE CHARACTERISTICS: 190 (A) LENGTH: 714 amino acids 191 (B) TYPE: amino acid 192 (C) STRANDEDNESS: single 193 (D) TOPOLOGY: linear 195 (Vii) IMMEDIATE SOURCE: 196 (A) LIBRARY: GenBank 197 (B) CLONE: 128842 199 (Xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3: 201 Met Val Lys Leu Ala Lys Ala Gly Lys Thr His Gly Glu Ala Lys Lys 202 1 5 10 15 203 Met Ala Pro Pro Pro Lys Glu Val Glu Glu Asp Ser Glu Asp Glu Glu 204 20 25 30 205 Met Ser Glu Glu Glu Asp Asp Ser Ser Gly Glu Glu Val Val Ile Pro 206 35 40 45 207 Gln Lys Lys Gly Lys Lys Ala Thr Ala Thr Pro Ala Lys Lys Val Val 208 50 55 60 209 Val Ser Gln Thr Lys Lys Val Ala Val Pro Thr Pro Ala Lys Lys Ala 210 65 70 70 70 75 80 211 Ala Val Thr Pro Gly Lys Lys Ala Ala Ala Thr Pro Gly Lys Lys Ala 212 85 90 95 213 Val Thr Pro Ala Lys Ala Val Ala Thr Pro Gly Lys Lys Gly Ala Thr 214 100 105 110 215 Gln Ala Lys Ala Leu Val Ala Thr Pro Gly Lys Lys Gly Ala Val Thr 216 115 120 125 217 Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp Ser																			
190 (A) LENGTH: 714 amino acids 191 (B) TYPE: amino acid 192 (C) STRANDEDNESS: single 193 (D) TOPOLOGY: linear 195 (vii) IMMEDIATE SOURCE: 196 (A) LIBRARY: GenBank 197 (B) CLONE: 128842 199 (Xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3: 201 Met Val Lys Leu Ala Lys Ala Gly Lys Thr His Gly Glu Ala Lys Lys 202 1 5 10 15 203 Met Ala Pro Pro Pro Lys Glu Val Glu Glu Asp Ser Glu Asp Glu Glu 204 20 25 30 205 Met Ser Glu Glu Glu Asp Asp Ser Ser Gly Glu Glu Val Val Ile Pro 206 35 40 45 207 Gln Lys Lys Gly Lys Lys Ala Thr Ala Thr Pro Ala Lys Lys Val Val 208 50 55 60 209 Val Ser Gln Thr Lys Lys Val Ala Val Pro Thr Pro Ala Lys Lys Ala 210 65 70 75 80 221 Ala Val Thr Pro Gly Lys Lys Ala Ala Ala Ala Thr Pro Ala Lys Lys Ala 222 85 90 95 233 Val Thr Pro Ala Lys Ala Val Ala Thr Pro Gly Lys Lys Gly Ala Thr 234 100 105 110 235 120 125 267 Pro Ala Lys Ala Leu Val Ala Thr Pro Gly Lys Lys Gly Ala Val Thr 236 115 120 125 268 179 Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp Ser	189																		
192		· ·																	
192	191	, ,																	
193 (Vii) IMMEDIATE SOURCE: 196 (A) LIBRARY: GenBank 197 (B) CLONE: 128842 199 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3: 201 Met Val Lys Leu Ala Lys Ala Gly Lys Thr His Gly Glu Ala Lys Lys 202 1 5 10 15 203 Met Ala Pro Pro Pro Lys Glu Val Glu Glu Asp Ser Glu Asp Glu Glu 204 20 25 30 205 Met Ser Glu Glu Glu Asp Asp Ser Ser Gly Glu Glu Val Val Ile Pro 206 35 40 45 207 Gln Lys Lys Gly Lys Lys Ala Thr Ala Thr Pro Ala Lys Lys Val Val 208 50 55 60 209 Val Ser Gln Thr Lys Lys Val Ala Val Pro Thr Pro Ala Lys Lys Ala 210 65 70 70 75 80 211 Ala Val Thr Pro Ala Lys Ala Val Ala Ala Ala Thr Pro Gly Lys Lys Ala 212 85 90 95 213 Val Thr Pro Ala Lys Ala Val Ala Thr Pro Gly Lys Lys Gly Ala Thr 214 100 105 110 215 Gln Ala Lys Ala Leu Val Ala Thr Pro Gly Lys Lys Gly Ala Val Thr 216 115 120 125 217 Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp Ser																			
196	193																		
197 (B) CLONE: 128842 199 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3: 201 Met Val Lys Leu Ala Lys Ala Gly Lys Thr His Gly Glu Ala Lys Lys 202 1 5 10 15 203 Met Ala Pro Pro Pro Lys Glu Val Glu Glu Asp Ser Glu Asp Glu Glu 204 20 25 30 205 Met Ser Glu Glu Glu Asp Asp Ser Ser Gly Glu Glu Val Val Ile Pro 206 35 40 45 207 Gln Lys Lys Gly Lys Lys Ala Thr Ala Thr Pro Ala Lys Lys Val Val 208 50 55 60 209 Val Ser Gln Thr Lys Lys Val Ala Val Pro Thr Pro Ala Lys Lys Ala 210 65 70 70 75 80 211 Ala Val Thr Pro Gly Lys Lys Ala Ala Ala Thr Pro Gly Lys Lys Ala 212 85 90 95 213 Val Thr Pro Ala Lys Ala Val Ala Thr Pro Gly Lys Lys Gly Ala Thr 214 100 105 110 215 Gln Ala Lys Ala Leu Val Ala Thr Pro Gly Lys Lys Gly Ala Val Thr 216 115 120 125 217 Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp Ser	195	, ,																	
199 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3: 201 Met Val Lys Leu Ala Lys Ala Gly Lys Thr His Gly Glu Ala Lys Lys 202 1	196	,	•	(A) LI	BRAR	Y: G	enBai	nk										
201 Met Val Lys Leu Ala Lys Ala Gly Lys Thr His Gly Glu Ala Lys Lys 202 1	197			(B) CL	ONE:	1288	342											
202	199	(xi)	SEQ	JENC:	E DES	SCRII	OLTS	N: SI	EQ II	ON C	: 3:							
203 Met Ala Pro Pro Pro Lys Glu Val Glu Glu Asp Ser Glu Asp Glu Glu 204	201	Met	Val	Lys	Leu	Ala	Lys	Ala	Gly	Lys	Thr	His	Gly	Glu	Ala	Lys	Lys		
204	202	1		_		5					10					15			
205 Met Ser Glu Glu Glu Asp Asp Ser Ser Gly Glu Glu Val Val Ile Pro 206	203	Met	Ala	Pro	Pro	Pro	Lys	Glu	Val	Glu	Glu	Asp	Ser	Glu	Asp	Glu	Glu		
206	204				20		_			25					30				
207 Gln Lys Lys Gly Lys Lys Ala Thr Ala Thr Pro Ala Lys Lys Val Val 208 50 55 60 209 Val Ser Gln Thr Lys Lys Val Ala Val Pro Thr Pro Ala Lys Lys Ala 210 65 70 70 75 80 211 Ala Val Thr Pro Gly Lys Lys Ala Ala Ala Thr Pro Ala Lys Lys Ala 212 85 90 95 213 Val Thr Pro Ala Lys Ala Val Ala Thr Pro Gly Lys Lys Gly Ala Thr 214 100 100 105 110 215 Gln Ala Lys Ala Leu Val Ala Thr Pro Gly Lys Lys Gly Ala Val Thr 216 115 120 125 217 Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp Ser	205	Met	Ser	Glu	Glu	Glu	Asp	Asp	Ser	Ser	Gly	Glu	Glu	Val	Val	Ile	Pro		
208	206			35			_	_	40					45					
209 Val Ser Gln Thr Lys Lys Val Ala Val Pro Thr Pro Ala Lys Lys Ala 210 65 70 75 80 211 Ala Val Thr Pro Gly Lys Lys Ala Ala Ala Thr Pro Ala Lys Lys Ala 212 85 90 95 213 Val Thr Pro Ala Lys Ala Val Ala Thr Pro Gly Lys Lys Gly Ala Thr 214 100 105 110 215 Gln Ala Lys Ala Leu Val Ala Thr Pro Gly Lys Lys Gly Ala Val Thr 216 115 120 125 217 Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp Ser	207	Gln	Lys	Lys	Gly	Lys	Lys	Ala	Thr	Ala	Thr	Pro	Ala	Lys	Lys	Val	Val		
210 65 70 75 80 211 Ala Val Thr Pro Gly Lys Lys Ala Ala Ala Thr Pro Ala Lys Lys Ala 212 85 90 95 213 Val Thr Pro Ala Lys Ala Val Ala Thr Pro Gly Lys Lys Gly Ala Thr 214 100 105 110 215 Gln Ala Lys Ala Leu Val Ala Thr Pro Gly Lys Lys Gly Ala Val Thr 216 115 120 125 217 Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp Ser	208			•	•	•	•							-	-				
210 65 70 75 80 211 Ala Val Thr Pro Gly Lys Lys Ala Ala Ala Thr Pro Ala Lys Lys Ala 212 85 90 95 213 Val Thr Pro Ala Lys Ala Val Ala Thr Pro Gly Lys Lys Gly Ala Thr 214 100 105 110 215 Gln Ala Lys Ala Leu Val Ala Thr Pro Gly Lys Lys Gly Ala Val Thr 216 115 120 125 217 Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp Ser	209	Val	Ser	Gln	Thr	Lys	Lys	Val	Ala	Val	Pro	Thr	Pro	Ala	Lys	Lys	Ala		
211 Ala Val Thr Pro Gly Lys Lys Ala Ala Ala Thr Pro Ala Lys Lys Ala 212 85 90 95 213 Val Thr Pro Ala Lys Ala Val Ala Thr Pro Gly Lys Lys Gly Ala Thr 214 100 105 110 215 Gln Ala Lys Ala Leu Val Ala Thr Pro Gly Lys Lys Gly Ala Val Thr 216 115 120 125 217 Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp Ser	210					-									_	_			
212 85 90 95 213 Val Thr Pro Ala Lys Ala Val Ala Thr Pro Gly Lys Lys Gly Ala Thr 214 100 105 110 215 Gln Ala Lys Ala Leu Val Ala Thr Pro Gly Lys Lys Gly Ala Val Thr 216 115 120 125 217 Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp Ser		Ala	Val	Thr	Pro	Gly	Lys	Lys	Ala	Ala	Ala	Thr	Pro	Ala	Lys	Lys	Ala		
214 100 105 110 215 Gln Ala Lys Ala Leu Val Ala Thr Pro Gly Lys Lys Gly Ala Val Thr 216 115 120 125 217 Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp Ser							•	-											
214 100 105 110 215 Gln Ala Lys Ala Leu Val Ala Thr Pro Gly Lys Lys Gly Ala Val Thr 216 115 120 125 217 Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp Ser		Val	Thr	Pro	Ala	Lys	Ala	Val	Ala	Thr	Pro	Gly	Lys	Lys	Gly	Ala	Thr		
215 Gln Ala Lys Ala Leu Val Ala Thr Pro Gly Lys Lys Gly Ala Val Thr 216 115 120 125 217 Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp Ser						-						-	,	-					
216 115 120 125 217 Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp Ser		Gln	Ala	Lys	Ala	Leu	Val	Ala	Thr		Gly	Lys	Lys	Gly	Ala	Val	Thr		
217 Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp Ser											-	-	-						
		Pro	Ala	Lys	Gly	Ala	Lys	Asn	Gly	Lys	Asn	Ala	Lys	Lys	Glu	Asp	Ser		
	218		130					135					140						

DATE: 02/04/2002

PATENT APPLICATION: US/09/978,242 TIME: 11:38:59

219 220	Asp 145	Glu	Asp	Glu	Asp	Asp 150	Asp	Asp	Asp	Glu	Asp 155	Asp	Ser	Asp	Glu	Asp 160
221	Glu	Glu	Asp	Glu			Asp	Glu	Phe		Pro	Pro	Val	Val	Lys 175	Gly
222				_	165	- 1	- 1 -	. 1 -	31-	170	. ז	Cor	Clu	\ cn		λcn
223 224	Lys	Gln	Gly	Lys 180	Val	Ala	Ala	Ala	185	Pro	Ala	ser	GIU	190	Gru	ASP
225	Glu	Glu	Glu		Glu	Glu	Glu	Glu	Glu	Glu	Asp	Glu	Glu	Glu	Glu	Asp
226	O.L.	02.0	195	1				200			_		205			
227	Asp	Ser		Glu	Glu	Glu	Ala	Met	Glu	Ile	Thr	Pro	Ala	Lys	Gly	Lys
228		210					215					220				
229	Lvs	Ala	Pro	Ala	Lys	Val	Val	Pro	Val	Lys	Ala	Lys	Asn	Val	Ala	Glu
230	225					230					235					240
231	Glu	Asp	Asp	Asp	Asp	Glu	Glu	Glu	Asp	Glu	Asp	Glu	Glu	Glu	Asp	Glu
232					245					250					255	
233	Glu	Glu	Glu	Glu	Asp	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu
234				260					265					270		
235	Pro	Val	Lys	Pro	Ala	Pro	Gly	Lys	Arg	Lys	Lys	Glu	Met	Thr	Lys	Gln
236			275					280					285			
237	Lys	Glu	Val	Pro	Glu	Ala	Lys	Lys	Gln	Lys	Val	Glu	Gly	Ser	Glu	Ser
238		290					295					300				
239	Thr	Thr	Pro	Phe	Asn	Leu	Phe	Ile	Gly	Asn		Asn	Pro	Asn	Lys	Ser
240	305					310					315		_		_	320
241	Val	Ala	Glu	Leu	Lys	Val	Ala	Ile	Ser		Pro	Phe	Ala	Lys	Asn	Asp
242					325					330			_	en 1	335	m
243	Leu	Ala	Val		Asp	Val	Arg	Thr		Thr	Asn	Arg	Lys	Phe	GIY	Tyr
244				340					345	_,	_	- 1	_	350	T	mba
245	Val	Asp		Glu	Ser	Ala	Glu	Asp	Leu	GLu	Lys	Ala		GIU	ьeu	1111
246			355	_				360	-1	T	T	<i>α</i> 1	365	Dro	Luc	C111
247	Gly		Lys	Val	Phe	GLy		GLu	TTE	Lys	Leu		ьуѕ	PIO	гуѕ	Gly
248		370					375			3	m L	380	Lou	λΙο	Tuc	λcn
249		Asp	Ser	Lys	Lys		Arg	Ala	Ala	Arg		Leu	ьeu	Ald	гуу	Asn 400
250	385	_	_ ,	_	~ 1	390	a1	3 ~~	<i>α</i> 1	T 011	395	Clu	17 - 1	Dha	Glu	
251	Leu	Ser	Phe	Asn		Thr	Glu	Asp	GIU	410	ьуѕ	GIU	vai	FILE	415	изъ
252			a1	T1.	405	Tau	17 - 1	Ser	Cln		Gly	Lve	Ser	Lvs		Tle
253	Ala	Leu	GIU	420	AIG	Leu	val	ser	425		GIY	цуз	501	430	O L 1	110
254	7 l a	Mirro	т1о		Dho	Tue	Sar	Glu			Ala	Glu	Lvs		Leu	Glu
255	Ald	1 7 1	435	GLU	rne	БуЗ	JCI	440	1114	пор	1114	014	445			
256	C111	Tirc	433 Cln	Clu	λla	Glu	Tle		Glv	Ara	Ser	Val			Tvr	Tyr
257 258	Glu	450		GIY	AIG	Giu	455		011	1119	001	460			1	1
259	Thr	430 C1v	Clu	Twe	Gly	Gln	Ara	Gln	Glu	Ara	Thr		Lvs	Asn	Ser	Thr
260	465		GIU	цуз	Gry	470	1119	0111	014	1.129	475		-1-			480
261			Gly	Glu	Ser		Thr	Leu	Val	Leu	Ser	Asn	Leu	Ser	Tyr	Ser
262	111	261	Oly	Olu	485				,	490					495	
263	Ala	Thr	Glu	Glu			Gln	Glu	Val	Phe	Glu	Lys	Ala	Thr	Phe	Ile
264		1111	<u> </u>	500					505			-		510		
265	Lvs	Va l	Pro	Gln	Asn	Gln	Gln	Gly			Lys	Gly	Tyr	Ala	Phe	Ile
266			515					520					525			
267	Glu	Phe	Ala	Ser	Phe	Glu	Asp	Ala	Lys	Glu	Ala	Leu	Asn	Ser	Cys	Asn
		_					-									

VERIFICATION SUMMARY

DATE: 02/04/2002

PATENT APPLICATION: US/09/978,242 TIME: 11:39:00

Input Set : N:\Crf3\RULE60\09978242.raw Output Set: N:\CRF3\02042002\I978242.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]